Uplift Modelling in Banking

Objective: Identify customers who are likely to be positively influenced by targeted marketing interventions, thereby driving incremental transactions and improving overall campaign effectiveness.

Key Modelling Algorithm

1. **Uplift Modelling:**

   - **Description:** Predicts the causal effect of a treatment (e.g., a targeted ad) on an individual's behavior. It aims to estimate the difference in outcomes between treated and untreated groups.

   - **Use Case:** Determine which customers are most likely to respond positively to marketing campaigns and tailor interventions accordingly.

Important Variables for Uplift Modelling (Indian Banking Perspective)

1. **Customer Profile:**

   - **Attributes:** Age, income, occupation, credit score (e.g., CIBIL score).

   - **Derived Attributes:** Risk profile, financial stability, spending patterns.

2. **Behavioral Patterns:**

   - **Attributes:** Purchase history, response to previous marketing campaigns, online activity.

   - **Derived Attributes:** Engagement score, response likelihood, churn propensity.

3. **Campaign Data:**

   - **Attributes:** Type of campaign, channel of communication, offer details, frequency of contact.

   - **Derived Attributes:** Campaign effectiveness score, historical uplift response.

Detailed Example of Uplift Modelling Implementation

Step-by-Step Implementation Using Uplift Modelling

1. **Data Preparation:**

   - Collect historical campaign data including both treated (those who received the marketing intervention) and control (those who did not receive the intervention) groups.

   - Preprocess data to handle missing values, encode categorical variables, and normalize numerical features.

2. **Feature Engineering:**

   - **Attributes:**

     - Age

     - Income

     - Occupation

     - Credit score (CIBIL score)

     - Purchase history

     - Response to previous marketing campaigns

     - Online activity

     - Type of campaign

     - Channel of communication

     - Offer details

     - Frequency of contact

   - **Derived Attributes:**

     - Risk profile

     - Financial stability

     - Spending patterns

     - Engagement score

     - Response likelihood

     - Churn propensity

     - Campaign effectiveness score

     - Historical uplift response

3. **Uplift Modelling Implementation:**

```python
import pandas as pd
from causalml.inference.tree import UpliftRandomForestCl
assifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report

# Sample data preparation
data = {
    'age': [35, 40, 45, 30, 50],
    'income': [60000, 80000, 120000, 50000, 110000],
    'occupation': ['salaried', 'business', 'salaried',
'student', 'retired'],
    'cibil_score': [750, 700, 800, 650, 780],
    'purchase_history': [5, 2, 4, 1, 3],
    'response_to_previous_campaigns': [1, 0, 1, 0, 1],
    'online_activity': [10, 5, 15, 3, 8],
    'campaign_type': ['email', 'sms', 'email', 'sms', 'e
mail'],
    'channel_of_communication': ['email', 'sms', 'emai
l', 'sms', 'email'],
    'offer_details': ['discount', 'cashback', 'discoun
t', 'cashback', 'discount'],
    'frequency_of_contact': [2, 1, 3, 1, 2],
    'treatment': [1, 0, 1, 0, 1],  # 1 for treated, 0 fo
r control
    'response': [1, 0, 1, 0, 1]  # 1 for positive respon
se, 0 for no response
}
df = pd.DataFrame(data)

# Split data into train and
```

```python
# Split data into train and test sets
X = df.drop(['response'], axis=1)
y = df['response']
```

```
    # Split into training and testing sets
    X_train, X_test, y_train, y_test = train_test_split(X,
y, test_size=0.2, random_state=42)

    # Train UpliftRandomForestClassifier
    uplift_model = UpliftRandomForestClassifier(n_estimators
=100, control_name='control')
    uplift_model.fit(X_train, treatment='treatment', y=y_tra
in)

    # Predict uplift on test set
    uplift_predictions = uplift_model.predict(X_test)

    # Evaluate the model
    print(classification_report(y_test, uplift_predictions))
```

1. **Interpretation and Actionable Insights:**

   - Analyze uplift scores to identify which customers are likely to be positively influenced by the campaign.

   - Focus marketing efforts on these high-uplift customers to maximize incremental transactions and campaign effectiveness.

   - Adjust marketing strategies based on the model's insights to reduce negative impacts on customers who may be adversely affected by the campaign.

Variables for Targeting

1. **Uplift Segment:**

   - **Attributes:**

     - Age

     - Income

     - Occupation

     - Credit score (CIBIL score)

     - Purchase history

- Response to previous marketing campaigns

  - Online activity

  - Type of campaign

  - Channel of communication

  - Offer details

  - Frequency of contact

- **Derived Attributes:**

  - Risk profile

  - Financial stability

  - Spending patterns

  - Engagement score

  - Response likelihood

  - Churn propensity

  - Campaign effectiveness score

  - Historical uplift response

- **Target Variables:**

  - Uplift score (likelihood of positive response to the campaign)

  - Incremental transaction likelihood

  - Expected increase in customer lifetime value (CLV)

Integrating Uplift Modelling with Look-alike Modelling

To further enhance campaign effectiveness, we can combine uplift modelling with look-alike modelling:

1. **Look-alike Modelling:**

   - Identify potential new customers who resemble existing high-value customers.

   - Use attributes like transaction history, product holdings, creditworthiness, and demographic information.

2. **Uplift Modelling:**

- Determine which of these look-alike customers are likely to respond positively to marketing interventions.

- Use uplift scores to target high-potential customers with personalized campaigns.

Implementation Flow

1. **Segmentation and Targeting:**

   - Segment customers based on demographic, behavioral, and transactional data.

   - Identify key segments such as students, high-profile customers, and seniors.

2. **Look-alike Modelling:**

   - Train a model on seed data (existing high-value customers) to find similar customers in the pool data.

   - Use attributes like age, income, occupation, credit score, transaction history, product holdings, and engagement levels.

3. **Uplift Modelling:**

   - Predict the causal effect of marketing interventions on customer behavior.

   - Target customers with high uplift scores to maximize incremental transactions and campaign ROI.

   - Use attributes like age, income, occupation, credit score, purchase history, response to previous campaigns, online activity, campaign type, channel of communication, offer details, and frequency of contact.

4. **Cross-Selling and Upselling:**

   - Implement recommendation systems to suggest relevant products to customers based on their transaction history and preferences.

   - Use attributes like transaction frequency, transaction amount, product usage, average transaction value, product affinity score, product usage trend, engagement score, response likelihood, and churn propensity.

By integrating these methodologies, the bank can create a comprehensive campaign strategy that effectively targets high-potential customers, maximizes

positive responses, and optimizes marketing spend.